

# THE PRODUCT OF MATRIX SUBSPACES

MARKO HUHTANEN\*

**Abstract.** In factoring matrices into the product of two matrices operations are typically performed with elements restricted to matrix subspaces. Such modest structural assumptions are realistic, for example, in large scale computations. This paper is concerned with analyzing associated matrix geometries. Curvature of the product of two matrix subspaces is assessed. As an analogue of the internal Zappa-Szép product of a group, the notion of factorizable matrix subspace arises. Interpreted in this way, several classical instances are encompassed by this structure. The Craig-Sakamoto theorem fits naturally into this framework.

**Key words.** operator factoring, product of matrix subspaces, curvature, irreducible matrix subspace

**AMS subject classifications.** 15A30, 47L05

**1. Introduction.** This paper is concerned with the set consisting of matrix products of elements restricted to matrix subspaces  $\mathcal{V}_1$  and  $\mathcal{V}_2$  of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ). Matrix subspaces appear regularly in large scale computational problems where only modest structural assumptions can be made.<sup>1</sup> For example, factorization problems [33, 24, 15] are typically subspace problems. The study of matrix subspaces can be classified as being finite dimensional operator space theory. For operator space theory, see [9, 29]. Regarding the geometry associated with matrix factoring [20, 4], the set of products is defined as

$$\mathcal{V}_1 \mathcal{V}_2 = \{V_1 V_2 : V_1 \in \mathcal{V}_1 \text{ and } V_2 \in \mathcal{V}_2\}.$$

As illustrated by the LU factorization, both complete and incomplete, as well as the singular value decomposition, this structure is ubiquitous. (The subset of  $\mathbb{C}^{n \times n}$  of matrices of rank  $k$  at most is also the product of two matrix subspaces of  $\mathbb{C}^{n \times n}$ .) These examples also underscore how different aims and geometries the set of products can have.

The set of products is not the most natural structure from the computational viewpoint of factoring.<sup>3</sup> It is encountered regularly, though. Through the LU factorization, band matrices is one such instance, even though it is exceptional by being a matrix subspace. In general, the set of products is constructible and certainly not flat. Thereby one is led to ask how curved  $\mathcal{V}_1 \mathcal{V}_2$  is. By locally inspecting the image of the smooth map

$$(V_1, V_2) \mapsto V_1 V_2, \tag{1.1}$$

this can be assessed Riemannian geometrically. This approach turns out to have a global character, leading to a necessary and sufficient condition for the flatness of the whole image of (1.1).

---

\* Department of Mathematics and Systems Analysis, Aalto University, Box 1100 FIN-02015, Finland, (Marko.Huhtanen@hut.fi). Supported by the Academy of Finland.

<sup>1</sup>In large scale problems, a matrix subspace typically defined by fixing a sparsity structure with  $O(n)$  nonzero entries.

<sup>2</sup>In algebraic geometry, when  $\mathcal{V}_1$  and  $\mathcal{V}_2$  are treated as projective Hilbert spaces, the tensor product is the more usual object of interest (given by the Segre map).

<sup>3</sup>Being computationally far more accessible, it is likely that the set  $\mathcal{V}_1 \text{Inv}(\mathcal{V}_2) = \{V_1 V_2^{-1} : V_1 \in \mathcal{V}_1 \text{ and } V_2 \in \mathcal{V}_2 \cap \text{GL}(n, \mathbb{C})\}$  is more important; see [20, 4].

Vanishing curvature is intriguing by the fact that then the associated matrix subspace factors, i.e., it is not a so-called irreducible matrix subspace. And conversely, it is a fundamental problem to recover whether a given matrix subspace is irreducible. Once interpreted this way, it becomes clear that this notion has many appearances already in the commutative case, such as integer factorization and polynomial factoring. In the noncommutative case of  $\dim \mathcal{V}_1 = \dim \mathcal{V}_2 = 2$ , an appropriate general treatment of the Craig-Sakamoto theorem in statistics is shown to correspond to such an instance.

The paper is organized as follows. In Section 2 the product of matrix subspaces is defined. Closedness is addressed and some fundamental linear algebraic notions are recalled. The structure is illustrated with several examples. Through the Craig-Sakamoto theorem, special attention is paid to the three dimensional case. Bihomogeneous polynomial maps of bidegree  $(1, 1)$  are associated with the problem. In Section 3 the notion of factorizable matrix subspace is introduced. The curvature of the set of products is assessed locally. Based on this, a necessary and sufficient condition for the flatness of the set of products is given.

**2. The product of two matrix subspaces.** The product of two matrix subspaces is defined and exemplified. Special attention is paid to a generalization of the Craig-Sakamoto theorem. Bihomogeneous polynomial maps are associated with the product of two matrix subspaces.

**2.1. Matrix subspaces and the Craig-Sakamoto theorem.** Assume  $\mathcal{V}$  is a matrix subspace of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ). Depending on whether  $\mathcal{V}$  contains invertible elements, the matrix subspace is called either nonsingular or singular [20]. Generically, a matrix subspace is nonsingular [22]. Then its subset consisting of invertible elements is open and dense [20].

Two matrix subspaces  $\mathcal{V}$  and  $\mathcal{W}$  are said to be equivalent if  $\mathcal{W} = X\mathcal{V}Y^{-1}$  holds for invertible matrices  $X, Y \in \mathbb{C}^{n \times n}$ . In view of their properties, equivalent matrix subspaces can in many ways be regarded as being indistinguishable.

**DEFINITION 2.1.** *Suppose  $\mathcal{V}_1$  and  $\mathcal{V}_2$  are matrix subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ). Then*

$$\mathcal{V}_1\mathcal{V}_2 = \{V_1V_2 : V_1 \in \mathcal{V}_1 \text{ and } V_2 \in \mathcal{V}_2\}$$

*is said to be the set of products of  $\mathcal{V}_1$  and  $\mathcal{V}_2$ .*

Clearly, we have a homogeneous set, i.e., there holds  $t\mathcal{V}_1\mathcal{V}_2 = \mathcal{V}_1\mathcal{V}_2$  for any nonzero scalar  $t$ . However, unlike a matrix subspace, the set of products need not be closed. (Consider, for example, the LU factorization.) Closedness is certainly of importance, e.g., in stable numerical computations.<sup>4</sup>

**THEOREM 2.2.** *Assume  $\mathcal{V}_1$  and  $\mathcal{V}_2$  are matrix subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ) such that  $V_1V_2 = 0$  if and only if either  $V_1 = 0$  or  $V_2 = 0$ . Then  $\mathcal{V}_1\mathcal{V}_2$  is closed.<sup>5</sup>*

*Proof.* Consider the matrix product

$$(V_1, V_2) \mapsto V_1V_2$$

with  $V_1 \in \mathcal{V}_1$  and  $V_2 \in \mathcal{V}_2$ . For any scalars  $s$  and  $t$  we have  $(sV_1, tV_2) \mapsto stV_1V_2$ . Thereby, because of the assumptions, we can regard the matrix product as a map

<sup>4</sup>It is well-known that the computation of an LU factorization is not a stable process unless one uses partial pivoting.

<sup>5</sup>A natural related problem is as follows. Assume a matrix subspace  $\mathcal{V}_1$  of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ) is given. Find a matrix subspace  $\mathcal{V}_2$  of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ) of the largest possible dimension such that the assumptions of Theorem 2.2 are satisfied.

from the product of projective spaces  $\mathbf{P}(\mathcal{V}_1) \times \mathbf{P}(\mathcal{V}_2)$  to the projective space  $\mathbf{P}(\mathbb{C}^{n \times n})$ . This is a map from a compact space to a compact space. Consequently, by the closed map lemma [27, Lemma 4.25], the image of  $\mathbf{P}(\mathcal{V}_1) \times \mathbf{P}(\mathcal{V}_2)$  is closed.  $\square$

For the computation of

$$\{(V_1, V_2) \in \mathcal{V}_1 \times \mathcal{V}_2 : V_1 V_2 = 0\},$$

see [10].

EXAMPLE 1. Let  $\mathcal{V}_1$  be the set of circulant matrices and  $\mathcal{V}_2$  the set of diagonal matrices in  $\mathbb{C}^{n \times n}$ . By Theorem 2.2, then  $\mathcal{V}_1 \mathcal{V}_2$  is closed. It is readily seen that now the matrix product corresponds to the so-called Segre map, a fundamental family of functions in algebraic geometry. For the Segre maps, see, e.g., [16, p.25]. (In other words, the Kronecker product of two vectors is really just an instance of the standard matrix product restricted to prescribed matrix subspaces.) The above theorem yields a natural generalization of such maps.

EXAMPLE 2. Certainly, the conditions of Theorem 2.2 are not necessary. For instance, the subset of  $\mathbb{C}^{n \times n}$  consisting of matrices of rank  $k$  at most is closed, a property of tremendous importance in approximating with the singular value decomposition. It equals the set of products  $\mathcal{V}_1 \mathcal{V}_2$ , where  $\mathcal{V}_1$  (resp.  $\mathcal{V}_2$ ) is the matrix subspace of  $\mathbb{C}^{n \times n}$  having the last  $n - k$  columns (resp. rows) zeros.

The following example supports the viewpoint that the set of products yields an equally natural “discretization” of Toeplitz operators as Toeplitz matrices do.

EXAMPLE 3. For an infinite dimensional noncommutative example, for invertible Toeplitz operators there is an upper-lower triangular factored Toeplitz structure; see, e.g., [25]. For Toeplitz matrices this product does not preserve Toeplitzness. The structure remains closed, though.

COROLLARY 2.3. *Denote by  $\mathcal{V}_1$  and  $\mathcal{V}_2$  the subspaces of upper and lower triangular Toeplitz matrices of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$ . Then  $\mathcal{V}_1 \mathcal{V}_2$  is closed.*

*Proof.* Consider the equation  $V_1 V_2 = 0$ . In the product, compute the last row first to have either the diagonal of  $V_1$  zero or  $V_2 = 0$ . Then proceed analogously upwards by computing next to the last row, to have the claim by using Theorem 2.2.  $\square$

The set of products in this corollary is intriguing since both subspaces are invertible.<sup>6</sup> Thereby it is a straightforward computation to recover if a given matrix  $A \in \mathbb{C}^{n \times n}$  belongs to  $\mathcal{V}_1 \mathcal{V}_2$  [20]. The closure of the set of inverses of invertible elements is the set of products  $\mathcal{V}_2 \mathcal{V}_1$ , which is closed by analogous arguments. Hence we have an elegant symmetry with respect to the inversion.<sup>7</sup>

Although not readily determined, the following quantity appears to be of central relevance. To the best of our knowledge, it was initially introduced in [3]. See also [11] for related computations.

DEFINITION 2.4. *Let  $\mathcal{V}$  be a matrix subspace of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ). Set*

$$\text{minrank}(\mathcal{V}) = \min_{V \in \mathcal{V}, V \neq 0} \text{rank}(V).$$

<sup>6</sup>Let  $\mathcal{V}$  and  $\mathcal{W}$  be two nonsingular matrix subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ). If

$$\{V^{-1} : V \in \mathcal{V} \cap \text{GL}(n, \mathbb{C})\} = \mathcal{W} \cap \text{GL}(n, \mathbb{C}),$$

then  $\mathcal{V}$  is said to be invertible [20].

<sup>7</sup>The problem of characterizing the inverses of Toeplitz matrices has been studied a lot; see, e.g., [12].

This quantity, let us call it the minrank of  $\mathcal{V}$ , is preserved under equivalence. In particular, in the nonsingular case of  $\dim \mathcal{V} = 2$ , the geometric multiplicities of the eigenvalues of a matrix determine  $\text{minrank}(\mathcal{V})$ . Namely, let  $\mathcal{V} = \text{span}\{V_1, V_2\}$  with  $V_1, V_2 \in \mathbb{C}^{n \times n}$ . Suppose  $X = I$  and  $Y$  is an invertible element of  $\mathcal{V}$ . Then

$$X\mathcal{V}Y^{-1} = \text{span}\{I, W_1\}, \quad (2.1)$$

so that it suffices to inspect the eigenvalues of  $W_1 \in \mathbb{C}^{n \times n}$  to determine the minrank of  $\mathcal{V}$ .

EXAMPLE 4. The so-called Hurwitz-Radon matrix subspace  $\mathcal{V}$  of  $\mathbb{C}^{n \times n}$  over  $\mathbb{R}$  has the property that any nonzero element is a scalar multiple of a unitary matrix. Hence, then we have  $\text{minrank}(\mathcal{V}) = n$ .

Using Proposition 2.2 we can conclude that, whenever

$$\text{minrank}(\mathcal{V}_1) + \text{minrank}(\mathcal{V}_2) > n,$$

then  $\mathcal{V}_1\mathcal{V}_2$  is closed.

The set of products need not be somehow curved by the fact that its closure can retain the structure of matrix subspace. Of course, we always have a subspace when either  $\mathcal{V}_1$  or  $\mathcal{V}_2$  is one dimensional.

EXAMPLE 5. For a matrix subspace admitting a “factorization” as the product of two matrix subspaces, consider the set of  $(p, q)$ -band matrices.<sup>8</sup> Denote by  $\mathcal{V}_1$  and  $\mathcal{V}_2$  the set of  $(p, 0)$ -band and  $(0, q)$ -band matrices. Then there exists a factorization of a nonsingular  $A \in \mathbb{C}^{n \times n}$  as  $A = V_1V_2$  if and only if  $A$  is a strongly nonsingular  $(p, q)$ -band matrix. As strong singularity is a generic property, we can conclude that the closure of  $\mathcal{V}_1\mathcal{V}_2$  is the set of  $(p, q)$ -band matrices, i.e., a matrix subspace.

Observe that between Examples 2 and 5 there is strong resemblance with regard to how the sparsity structures of the matrix subspaces are defined.

In statistics, the so-called Craig-Sakamoto theorem<sup>9</sup> is similarly related with a factorization of a matrix subspace; for details on the Craig-Sakamoto theorem, see [8]. Besides statistics, this structure is of interest in studying the spectrum in the case of three dimensional matrix subspaces [32]. It is, in essence, concerned with the claim of the following proposition in the case of  $k = 2$  and  $\dim \mathcal{V}_1 = \dim \mathcal{V}_2 = 1$ .

PROPOSITION 2.5. *Assume a matrix subspace  $\mathcal{V}$  of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  can be decomposed as  $\mathcal{V} = \mathbb{C}I + \sum_{j=1}^k \mathcal{X}_j$  with matrix subspaces  $\mathcal{X}_j$  satisfying  $\mathcal{X}_j\mathcal{X}_l = \{0\}$  for  $j < l$ . Then  $\mathcal{V}$  equals the closure of  $\prod_{j=1}^k (\mathbb{C}I + \mathcal{X}_j)$ .*

*Proof.* Assume  $V = tI + X_1 + \cdots + X_k \in \mathcal{V}$  with  $X_j \in \mathcal{X}_j$  and a scalar  $t \neq 0$ . (If  $t = 0$ , then  $V$  can be approximated with such an element.) By the fact that  $X_jX_l = 0$  for  $j < l$ , we can write  $V = (tI + X_1)(I + X_2/t) \cdots (I + X_k/t)$ .  $\square$

The appearing condition on the matrix subspaces  $\mathcal{X}_j$  forces them to be singular. It is noteworthy that this structure is also utilized in computing (inverting) the  $L$  factor in the LU decomposition of a matrix.

Let  $\mathcal{V}_1$  and  $\mathcal{V}_2$  be nonsingular matrix subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  of dimension two. (Due to the generalized eigenvalue problem, the two dimensional case is possibly the most frequently encountered.) In view of the Craig-Sakamoto theorem, it is natural to

<sup>8</sup>A  $(p, q)$ -band matrix is a square matrix with lower bandwidth  $p$  and upper bandwidth  $q$ .

<sup>9</sup>The Craig-Sakamoto theorem: Two real symmetric matrices  $X_1$  and  $X_2$  satisfy  $\det(I - tX_1 - sX_2) = \det(I - tX_1)\det(I - sX_2)$  for all  $t, s \in \mathbb{R}$  if and only if  $X_1X_2 = 0$ .

ask when the closure of  $\mathcal{V}_1\mathcal{V}_2$  is a matrix subspace. Here we may equally well consider equivalent matrix subspaces of the form

$$X\mathcal{V}_1 = \text{span}\{I, X_1\} \text{ and } \mathcal{V}_2Y^{-1} = \text{span}\{I, X_2\} \quad (2.2)$$

with invertible matrices  $X, Y \in \mathbb{C}^{n \times n}$  chosen appropriately. Then, instead of  $\mathcal{V}_1\mathcal{V}_2$ , it suffices to inspect the set of products  $X\mathcal{V}_1\mathcal{V}_2Y^{-1}$ .

**THEOREM 2.6.** *Let  $\mathcal{V}_1$  and  $\mathcal{V}_2$  be nonsingular matrix subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  of dimension 2. If the closure of  $\mathcal{V}_1\mathcal{V}_2$  is a matrix subspace, then it is of dimension 3 at most. This holds if and only if the matrices in (2.2) satisfy*

$$X_1(cX_2 - dI) = (aX_2 - bI) \quad (2.3)$$

for some constants  $a, b, c, d \in \mathbb{C}$  not all zero.

*Proof.* Suppose  $\text{span}\{V_1, V_2\} = \mathcal{V}_1$  and  $\text{span}\{V_3, V_4\} = \mathcal{V}_2$ . For the claim we may consider the map

$$(z_1, z_2, z_3, z_4) \mapsto z_1z_3V_1V_3 + z_1z_4V_1V_4 + z_2z_3V_2V_3 + z_2z_4V_2V_4 \quad (2.4)$$

from  $\mathbb{C}^4$  to  $\mathbb{C}^{n \times n}$ . If the closure of the image were four dimensional, then the set of equations

$$z_1z_3 = b_1, z_1z_4 = b_2, z_2z_3 = b_3 \text{ and } z_2z_4 = b_4 \quad (2.5)$$

should have a solution for a dense subset of vectors  $(b_1, b_2, b_3, b_4)$  of  $\mathbb{C}^4$ . Choose  $b_1 = t + \epsilon_1$ ,  $b_2 = t + \epsilon_2$ ,  $b_3 = -t + \epsilon_3$  and  $b_4 = t + \epsilon_4$  with  $t \in \mathbb{C} \setminus \{0\}$  and small  $|\epsilon_j| \ll |t|$ , for  $j = 1, \dots, 4$ . (For  $\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon_4 = 0$  we clearly have no solutions.) We obtain

$$z_3 = (1 + \hat{\epsilon})z_4 \text{ and } z_3 = (-1 + \tilde{\epsilon})z_4,$$

where  $\hat{\epsilon} = \frac{\epsilon_1 - \epsilon_2}{t + \epsilon_2}$  and  $\tilde{\epsilon} = \frac{\epsilon_3 - \epsilon_4}{t + \epsilon_4}$ . Subtracting leads to  $0 = (2 + \tilde{\epsilon} - \hat{\epsilon})z_4$  which is a contradiction for small small  $\epsilon_j$ . Hence there are no solutions near the point  $(1, 1, -1, 1)$ .

Assume the closure of  $\mathcal{V}_1\mathcal{V}_2$  is a matrix subspace. Then the matrices  $V_1V_3$ ,  $V_1V_4$ ,  $V_2V_3$  and  $V_2V_4$  are contained in this subspace. (For example, the choices  $z_1 = z_3 = 1$  and  $z_2 = z_4 = 0$  yield  $V_1V_2$ , etc.) If the dimension is 3, it follows that  $V_1V_3$ ,  $V_1V_4$ ,  $V_2V_3$  and  $V_2V_4$  are linearly dependent, i.e.,  $X_1(X_2 - dI) = aX_2 - bI$ . If the dimension is 2, then  $X_1 = aX_2 - bI$ .

For the converse, assume  $\dim\{V_1V_3, V_1V_4, V_2V_3, V_2V_4\} = 3$ . (The case of dimension 2 is trivial.) We may assume, let us say,  $V_1V_3 = \alpha_1V_1V_4 + \alpha_2V_2V_3 + \alpha_3V_2V_4$  with fixed  $\alpha_j \in \mathbb{C}$ , for  $j = 1, 2, 3$ . Inserting this into (2.4) gives

$$(\alpha_1z_1z_3 + z_1z_4)V_1V_4 + (\alpha_2z_1z_3 + z_2z_3)V_2V_3 + (\alpha_3z_1z_3 + z_2z_4)V_2V_4.$$

Let a vector  $b = (b_1, b_2, b_3)$  be given. After possibly an arbitrarily small perturbation, we may assume  $b_2 \neq 0$ . Then necessarily  $z_3 \neq 0$ , so that with  $w = \frac{z_4}{z_3}$  the problem of attaining these coefficients is equivalent to finding a solution to the equations

$$z_1(\alpha_1 + w) = c_1, \alpha_2z_1 + z_2 = c_2, \alpha_3z_1 + wz_2 = c_3 \quad (2.6)$$

with  $c_2 \neq 0$ . The last two equations, after eliminating  $z_2$ , can be combined into a single one  $z_1(\alpha_3 - \alpha_2w) + c_2w = c_3$ . This combined with the first equation in (2.6) results in

$$c_2w^2 + (c_2\alpha_1 - c_1\alpha_2 - c_3)w + c_2\alpha_3 - c_3\alpha_1 = 0$$

which should have a nonzero solution satisfying  $w \neq -\alpha_1$ . This can be achieved, after possibly an arbitrarily small perturbations of the coefficients  $c_1$ ,  $c_2$  and  $c_3$ .  $\square$

In particular, the Craig-Sakamoto theorem corresponds to the special case in which the constants satisfy  $a = b = d = 0$  and  $c = 1$ .

In the identity (2.3) we are dealing with what we call a generalized linear fractional transformation. Namely, if  $d/c$  is not an eigenvalue of  $X_2$ , then  $X_1$  is a classical linear fractional transformation of  $X_2$ . Combined with the equivalence transformations (2.2), this yields a way to construct two dimensional matrix subspaces such that the closure of the set of products is a matrix subspace.

**2.2. The set of products and bihomogeneous polynomial maps of bidegree  $(1, 1)$ .** There is a family of polynomial functions, studied especially in algebraic geometry, which is intimately connected with the product of two matrix subspaces. To describe this connection, for the set of products let us introduce its linearization defined as follows.

**DEFINITION 2.7.** Assume  $\mathcal{V}_1$  and  $\mathcal{V}_2$  are matrix subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ). The matrix subspace of the smallest possible dimension including  $\mathcal{V}_1 \mathcal{V}_2$  is said to be the linearization of  $\mathcal{V}_1 \mathcal{V}_2$ .

Assume  $V^1, \dots, V^j$  and  $V^{j+1}, \dots, V^k$  are bases of  $\mathcal{V}_1$  and  $\mathcal{V}_2$ , respectively. If  $W^1, \dots, W^l$  is a basis of the linearization  $\mathcal{W}$ , then the inclusion relation of the linearization implies that

$$V^s V^t = \sum_{r=1}^l m_r^{st} W^r$$

for some constants  $m_r^{st} \in \mathbb{C}$ . Thereby the product of arbitrary elements  $V_1 \in \mathcal{V}_1$  and  $V_2 \in \mathcal{V}_2$  can be written as

$$V_1 V_2 = \sum_{s=1}^j z_s V^s \sum_{t=j+1}^k w_t V^t = \sum_{r=1}^l z^T M_r w W^r \quad (2.7)$$

with  $z = (z_1, \dots, z_j) \in \mathbb{C}^j$ ,  $w = (w_{j+1}, \dots, w_k) \in \mathbb{C}^{k-j}$  and matrices  $M_r = \{m_r^{st}\} \in \mathbb{C}^{j \times (k-j)}$ . In terms of this expansion, the problem converts into inspecting the bihomogeneous polynomial map

$$(z, w) \mapsto M(z)w = \begin{bmatrix} z^T M_1 \\ \vdots \\ z^T M_l \end{bmatrix} w \quad (2.8)$$

of bidegree  $(1, 1)$  from  $\mathbb{C}^j \times \mathbb{C}^{k-j}$  to  $\mathbb{C}^l$ . (For computational algebraic geometric aspects of such functions, see [10].) Clearly,  $M : \mathbb{C}^j \rightarrow \mathbb{C}^{l \times (k-j)}$  is linear with the  $p$ th column equaling  $z^T M_p$  at  $z \in \mathbb{C}^j$ .

**EXAMPLE 6.** In the proof of Theorem 2.6, when the closure of  $\mathcal{V}_1 \mathcal{V}_2$  is the three dimensional matrix subspace  $\mathcal{W}$ , we have  $j = 2$ ,  $k = 4$  and  $l = 3$ . Then  $W = V_1 V_2$ ,  $W_2 = V_2 V_3$  and  $W = V_2 V_4$  with

$$M_1 = \begin{bmatrix} \alpha_1 & 1 \\ 0 & 0 \end{bmatrix}, M_2 = \begin{bmatrix} \alpha_2 & 0 \\ 1 & 0 \end{bmatrix} \text{ and } M_3 = \begin{bmatrix} \alpha_3 & 0 \\ 0 & 1 \end{bmatrix}.$$

The following theorem is of importance.

**THEOREM 2.8.** Let  $\mathcal{V}_1$ ,  $\mathcal{V}_2$  and  $\mathcal{W}$  be matrix subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$ . If the closure of  $\mathcal{V}_1 \mathcal{V}_2$  is  $\mathcal{W}$ , then  $\mathcal{V}_1 \mathcal{V}_2$  contains an open dense subset of  $\mathcal{W}$ .

*Proof.* Using arguments from algebraic geometry, this follows directly from a theorem of Chevalley, i.e., from the fact that the map (2.8) is regular and that its image is therefore constructible. (For more details, see, e.g., [28, Theorem 10.2] or [17, p. 94].) If the closure of the image (in the standard topology of  $\mathbb{C}^l$ ) is  $\mathbb{C}^l$ , then  $\mathbb{C}^l$  is also the Zariski closure of the image. Thus, it contains a Zariski open set of  $\mathbb{C}^l$ . Such a set is open and dense in  $\mathbb{C}^l$ .  $\square$

In principle, with (2.8) the problem of recovering whether a given matrix  $A \in \mathcal{W}$  is factorizable can be approached with the Hilbert Nullstellensatz. As opposed to considering (3.4), this problem belongs to the realm of commutative algebra. For the existence of a solution, by invoking the effective Nullstellensatz we have the linear algebra problem of solving a linear system. It is, however, only formally so because of the exponential growth of the size of the linear systems; see Appendix.

The converse problem is actually of interest. Namely, consider solving a bilinear system

$$M(z)w = b \quad (2.9)$$

with a given  $b \in \mathbb{C}^l$ . If an expansion (2.7) can be established, for some matrix subspaces  $\mathcal{V}_1$ ,  $\mathcal{V}_2$  and  $\mathcal{W}$ , with either  $\mathcal{V}_1$  or  $\mathcal{V}_2$  invertible then solving (2.9) is straightforward with the algorithms proposed in [20]. Consequently, the problem of finding such matrix subspaces for a given bilinear system (2.9) seems to be of central relevance.

**3. Riemannian geometry of the product of matrix subspaces.** Next the geometry of the set of products is inspected. The flat case is introduced first. Then the general case is studied in terms of smooth maps.

**3.1. Factorizable matrix subspaces.** The structure appearing in connection with the LU decomposition and the Craig-Sakamoto theorem is really just an instance of the following general notion of noncommutative factoring.

**DEFINITION 3.1.** *A matrix subspace  $\mathcal{W}$  of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ) is said to be factorizable if*

$$\mathcal{W} = \overline{\mathcal{V}_1 \mathcal{V}_2} \quad (3.1)$$

for matrix subspaces  $\mathcal{V}_1$  and  $\mathcal{V}_2$  of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ) satisfying the conditions

$$1 < \min\{\dim \mathcal{V}_1, \dim \mathcal{V}_2\} \quad \text{and} \quad \max\{\dim \mathcal{V}_1, \dim \mathcal{V}_2\} < \dim \mathcal{W}. \quad (3.2)$$

A matrix subspace which is not factorizable is said to be irreducible. As Example 5 illustrates, taking the closure may be necessary. Observe though that, for matrix subspaces over  $\mathbb{C}$ , the set of products  $\mathcal{V}_1 \mathcal{V}_2$  is topologically and measure theoretically large in  $\mathcal{W}$  by Theorem 2.8.

The structure is preserved under equivalence, i.e.,  $\mathcal{W}$  is factorizable if and only if  $X\mathcal{W}Y^{-1}$  is factorizable for any invertible  $X, Y \in \mathbb{C}^{n \times n}$ .

A central problem, typically tough for one reason or another, is to recover whether a given matrix subspace is factorizable. Needless to say, if a matrix subspace  $\mathcal{W}$  can be factored, the interest turns completely on the factors  $\mathcal{V}_1$  and  $\mathcal{V}_2$ .<sup>10</sup>

**EXAMPLE 7.** This is Example 3 continued. In finite dimensions, an analogous problem corresponds to asking if the set of Toeplitz matrices is a factorizable matrix

---

<sup>10</sup>This is manifested by the LU factorization. After computing an LU factorization, the original matrix is typically thrown away and only the factors are saved in the storage.

subspace. (It is likely of use to recall that Toeplitz matrices are equivalent to Hankel matrices.)

It is noteworthy that when the set of products is closed, the singular elements of  $\mathcal{W}$  are exactly determined by the singular elements of the factors  $\mathcal{V}_1$  and  $\mathcal{V}_2$ . (Otherwise only an inclusion can be guaranteed.) Hence, then the study of the spectrum of  $\mathcal{W}$ , i.e., its singular elements, reduces to the study of the spectra of  $\mathcal{V}_1$  and  $\mathcal{V}_2$ . This is precisely the case in the Craig-Sakamoto theorem.

EXAMPLE 8. Interpreting the spectrum algebraic geometrically, in the case of  $\mathcal{W} = \mathcal{V}_1 \mathcal{V}_2$  the determinantal variety of  $\mathcal{W}$  is determined by the determinantal varieties of  $\mathcal{V}_1$  and  $\mathcal{V}_2$ . This “factorization” is intriguing since the determinantal variety of  $\mathcal{W}$  may well be irreducible. For example, we have  $\mathbb{C}^{n \times n} = \mathcal{V}_1 \mathcal{V}_2$ , where  $\mathcal{V}_1 = \mathcal{V}_2$  is the set of symmetric matrices. (This is a classical result; see, e.g., [15] and [24] and references therein.)

In general, (3.1) is a noncommutative notion with respect to the factors. (Curiously, when  $\dim \mathcal{V}_1 = \dim \mathcal{V}_2 = 1$ , the commutative case corresponds to the  $\omega$ -commutativity of matrices [18].) Although more stringent, in group theory there is the so-called internal Zappa-Szép product of a group having some aspects in common with Definition 3.1.<sup>11</sup> For the internal Zappa-Szép product, see the paper [1], which is partially expository, and references therein.

Only the case where the conditions (3.2) are satisfied is of interest since otherwise we are dealing with the equivalence of matrix subspaces. Consequently, a two dimensional matrix subspace is never factorizable. Then, through the equivalence (2.1), the nonsingular case is completely understood in terms of canonical forms for matrices. Two dimensional matrix subspaces are hence fundamentally different and correspond, in essence, to classical matrix analysis. Observe that the first nontrivial case, i.e., the three dimensional case over  $\mathbb{C}$  can be regarded as understood by Theorem 2.6.

In Example 2 there is the growth of  $k$  and in Example 5 the growth of the bandwidth, i.e., we have natural hierarchies of sets of products. The maximum values of the indices correspond to the set of products  $\mathbb{C}^{n \times n}$ . Analogously, suppose (3.1) holds. Take two sequences of nested subspaces

$$\mathcal{V}_1^1 \subset \mathcal{V}_1^2 \subset \dots \subset \mathcal{V}_1^{\dim \mathcal{V}_1} = \mathcal{V}_1 \quad \text{and} \quad \mathcal{V}_2^1 \subset \mathcal{V}_2^2 \subset \dots \subset \mathcal{V}_2^{\dim \mathcal{V}_2} = \mathcal{V}_2.$$

Then  $\{\mathcal{V}_1^j \mathcal{V}_2^k\}_{j,k}$  yields a natural hierarchy of sets of products such that for the maximum values of the indices the closure of the set of products is  $\mathcal{W}$ .

After these general remarks, let us give further illustrations on how Definition 3.1 actually encompasses numerous classical instances.

EXAMPLE 9. Although we are concerned with finite dimensional (noncommutative matrix) subspaces, infinite dimensional (commutative) problems are certainly of equal interest. For the classical problem of integer factoring, let  $k \in \mathbb{Z}$  be given. Associate with  $k$  the set  $k\mathbb{Z}$ , regarded as subspace of  $\mathbb{Z}$  over  $\mathbb{Z}$ . Then the question of whether  $k$  is a irreducible is equivalent to asking whether  $k\mathbb{Z}$  is a factorizable subspace.

For another commutative case, consider the ring of complex polynomials. To this corresponds a very classical notion. Namely, denote by  $\mathcal{V}_k$  the subspace consisting of complex polynomials of degree  $k$  at most. (Of course, factoring depends heavily

---

<sup>11</sup>Let  $G$  be a group with the identity element  $e$ , and let  $H$  and  $K$  be subgroups of  $G$ . If  $G = HK$  and  $H \cap K = \{e\}$ , then  $G$  is the internal Zappa-Szép product of  $H$  and  $K$ .



on the choice of field.) Then the fundamental theorem of algebra can be stated as a factorization result

$$\mathcal{V}_k = \mathcal{V}_1 \mathcal{V}_{k-1}$$

for  $\mathcal{V}_k$ . (This can also be formulated as an operator factorization result in infinite dimensions for lower triangular Toeplitz operators with finite bandwidth.) Of course, this is a tremendously powerful fact with practical implications, i.e., one would certainly like to have any given polynomial factored.

EXAMPLE 10. Polynomial factoring relates to matrix subspaces as follows. Suppose  $A \in \mathbb{C}^{n \times n}$  and consider

$$\mathcal{V} = \mathcal{K}(A; I) = \text{span}_{\mathbb{C}} \{I, A, A^2, \dots\}.$$

Then, by the fundamental theorem of algebra,  $\mathcal{V}$  is a factorizable matrix subspace whenever  $\deg(A) > 2$ .

**3.2. Geometry of the product of matrix subspaces.** In general, the set of products of two matrix subspaces  $\mathcal{V}_1$  and  $\mathcal{V}_2$  of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ) cannot be expected to be flat, i.e., to yield a factorizable matrix subspace. The study of its geometry involves linear structures, though. To this end, associate with the set of products the bilinear map

$$\Psi(V_1, V_2) = V_1 V_2 \tag{3.3}$$

from the direct sum  $\mathcal{V}_1 \times \mathcal{V}_2$  to  $\mathbb{C}^{n \times n}$ . Certainly,  $\Psi$  is smooth as it can be treated as a bihomogeneous polynomial map once a basis of the vector space  $\mathcal{V}_1 \times \mathcal{V}_2$  has been fixed appropriately; see (2.8).

EXAMPLE 11. Although not done in this paper, let us emphasize that it is also natural to study  $\Psi$  on subspaces of  $\mathcal{V}_1 \times \mathcal{V}_2$ . For example, take  $\mathcal{V}_1 = \mathcal{V}_2^T$  to be the set of lower triangular matrices in  $\mathbb{C}^{n \times n}$ . Then, related with the LU factoring of symmetric matrices, consider  $\Psi$  on the subspace  $\{(V_1, V_2) \in \mathcal{V}_1 \times \mathcal{V}_2 : V_1 = V_2^T\}$ .

If the linearization of  $\mathcal{V}_1 \mathcal{V}_2$  is not the whole  $\mathbb{C}^{n \times n}$ , then the set of products is said to be degenerate. In view of this, for matrix subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$ , a generalization of Picard's theorem [14] can be used to bound the dimension of the linearization. For this, assume  $A \in \mathbb{C}^{n \times n}$  and look at the map  $A - \Psi$  defined as

$$(V_1, V_2) \mapsto A - V_1 V_2 \tag{3.4}$$

If  $A \notin \mathcal{V}_1 \mathcal{V}_2$ , then this does not have zeros and hence we may apply the following theorem with  $m = \dim \mathcal{V}_1 + \dim \mathcal{V}_2$  and  $l = n^2 - 1$ .

THEOREM 3.2. [14] *Let  $f : \mathbb{C}^m \rightarrow \mathbb{P}_l(\mathbb{C})$  be a holomorphic map that omits  $l + k$  hyperplanes<sup>12</sup> in general position,  $k \geq 1$ . Then the image of  $f$  is contained in a projective linear subspace of dimension  $\leq [l/k]$ , where the brackets mean greatest integer.*

To inspect the structure of the set of products more locally, on any matrix subspace  $\mathcal{V}$  of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  the standard inner product

$$(V_1, V_2) = \text{tr}(V_2^* V_1) \text{ with } V_1, V_2 \in \mathcal{V} \tag{3.5}$$

---

<sup>12</sup>A hyperplane in homogeneous coordinates  $z_0, \dots, z_l$  is  $\{(z_0, \dots, z_l) : \sum_{j=0}^l a_j z_j = 0\}$  for fixed  $a_j \in \mathbb{C}$ .

is used. (Take the real part for matrix subspaces over  $\mathbb{R}$ .) The respective norm  $\|\cdot\|_F$  is the Frobenius norm. To proceed Riemannian geometrically, take a smooth curve with the Taylor expansion

$$c(t) = (V_1, V_2) + t(W_1, W_2) + t^2(U_1, U_2) + \dots$$

in  $\mathcal{V}_1 \times \mathcal{V}_2$  passing through a point  $(V_1, V_2)$ . Then  $\Psi$  maps this curve to  $\mathbb{C}^{n \times n}$  as

$$V_1 V_2 + t(V_1 W_2 + W_1 V_2) + t^2(V_1 U_2 + W_1 W_2 + U_1 V_2) + \dots \quad (3.6)$$

Considering its linearization, the linear terms span the matrix subspace

$$V_1 \mathcal{V}_2 + \mathcal{V}_1 V_2. \quad (3.7)$$

In particular, the dimension of this subspace yields the rank of  $\Psi$  at  $(V_1, V_2)$ .

It is clear that the maximum rank yields a lower bound on the dimension of the linearization of  $\mathcal{V}_1 \mathcal{V}_2$ . (Proof: (3.7) is a subset of the linearization of  $\mathcal{V}_1 \mathcal{V}_2$ .)

**PROPOSITION 3.3.** *Let  $\mathcal{V}_1$  and  $\mathcal{V}_2$  be two subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  (or  $\mathbb{R}$ ). Then  $\overline{\mathcal{V}_1 \mathcal{V}_2}$  equals its linearization if and only if  $\overline{\mathcal{V}_1 \mathcal{V}_2}$  contains the matrix subspace (3.7) for any  $(V_1, V_2)$ .*

*Proof.* Suppose  $\overline{\mathcal{V}_1 \mathcal{V}_2}$  is a subspace, i.e., equals its linearization. Since  $\Psi : \mathcal{V}_1 \times \mathcal{V}_2 \rightarrow \overline{\mathcal{V}_1 \mathcal{V}_2}$ , it follows that  $\overline{\mathcal{V}_1 \mathcal{V}_2}$  contains the matrix subspaces (3.7).

Suppose  $\overline{\mathcal{V}_1 \mathcal{V}_2}$  contains the matrix subspaces (3.7). Then  $V_1 Y + X V_2 \in \overline{\mathcal{V}_1 \mathcal{V}_2}$  for any  $V_1, X \in \mathcal{V}_1$  and  $Y, V_2 \in \mathcal{V}_2$ , i.e., it contains the sums. The multiples are contained by the homogeneity. Thereby  $\overline{\mathcal{V}_1 \mathcal{V}_2}$  is a subspace.  $\square$

The following facts are well known; see, e.g., [5, Chapter 2].

**PROPOSITION 3.4.** *The rank of  $\Psi$  generically attains the maximum.*

*Proof.* Consider the derivative of  $\Psi$  regarded as a function of  $m = \dim \mathcal{V}_1 + \dim \mathcal{V}_2$  variables after fixing bases of  $\mathcal{V}_1$  and  $\mathcal{V}_2$ . Recall that a matrix has rank  $r$  if and only if there exists at least one non-zero  $r$ -by- $r$  minor. Computing the determinant of the respective minor of the derivative gives a nonzero polynomial in  $m$  variables. The points where the maximum of the rank is not attained belong to its zero set.  $\square$

The following also justifies calling (3.7) locally the tangent space of  $\mathcal{V}_1 \mathcal{V}_2$  at  $\Psi(V_1, V_2) = V_1 V_2$ .

**PROPOSITION 3.5.** *Let the rank of  $\Psi$  attain the maximum at  $(V_1, V_2)$ . Then the image of a neighbourhood of  $(V_1, V_2)$  under  $\Psi$  is a smooth submanifold of  $\mathbb{C}^{n \times n}$  of the dimension equaling the rank.*

*Proof.* Let  $m = \dim \mathcal{V}_1 + \dim \mathcal{V}_2$  and denote by  $k$  rank of  $\Psi$  at  $(V_1, V_2)$ . Because of the constant rank theorem, we have the normal form

$$\psi \circ \Psi \circ \phi^{-1}(x_1, \dots, x_m) \mapsto (x_1, \dots, x_k, 0, 0, \dots, 0)$$

for appropriate charts  $\phi$  and  $\psi$ . Denote the components of  $\psi$  by  $\psi_j$  and consider the map  $(\psi_{k+1}, \dots, \psi_m)$  from  $\mathbb{C}^{n \times n}$  to  $\mathbb{C}^{n^2-k}$ . For this map, look at the inverse image of  $(0, 0, \dots, 0)$  to have the claim (a consequence of the constant rank theorem).  $\square$

Assume the rank of  $\Psi$  attains its maximum at  $(V_1, V_2)$ . By the above proposition, locally the image can be regarded as a smooth submanifold of  $\mathbb{C}^{n \times n}$ . With respect to the standard inner product on  $\mathbb{C}^{n \times n}$ , denote by  $\mathbf{P}_{V_1, V_2}$  the orthogonal projector on  $\mathbb{C}^{n \times n}$  onto the tangent space  $V_1 \mathcal{V}_2 + \mathcal{V}_1 V_2$ . Then to Riemannian geometrically assess how curved  $\mathcal{V}_1 \mathcal{V}_2$  is at  $\Psi(V_1, V_2)$ , consider the map

$$Q_{(V_1, V_2)}(W_1, W_2) = 2(I - \mathbf{P}_{V_1, V_2})W_1 W_2 \quad (3.8)$$

on  $\mathcal{V}_1 \times \mathcal{V}_2$ . Using this notation, its introduction can be argued as follows.

**PROPOSITION 3.6.** *Let the rank of  $\Psi$  attain its maximum at  $(V_1, V_2)$ . Then, for the image of a neighbourhood of  $(V_1, V_2)$  under  $\Psi$ , the extrinsic curvature of the geodesic passing through  $\Psi(V_1, V_2)$  with the speed vector  $V_1 W_2 + W_1 V_2$  equals*

$$\|Q_{(V_1, V_2)}(W_1, W_2)\|_F. \quad (3.9)$$

*Proof.* An application of (3.8) always corresponds to a curve (3.6) having the acceleration orthogonal to  $V_1 \mathcal{V}_2 + \mathcal{V}_1 V_2$  which is achieved by choosing the coefficient  $(U_1, U_2)$  such that  $V_1 U_2 + W_1 W_2 + U_1 V_2$  belongs to the orthogonal complement of the matrix subspace  $V_1 \mathcal{V}_2 + \mathcal{V}_1 V_2$ . Such a property is required from the geodesics passing through  $\Psi(V_1, V_2)$ ; see [26, pp. 138–139]. Hence, for a geodesics passing through  $\Psi(V_1, V_2)$  with the speed vector  $V_1 W_2 + W_1 V_2$ , the coefficient  $(U_1, U_2)$  is determined by this condition.  $\square$

If the map (3.8) vanishes identically, then the set of products  $\mathcal{V}_1 \mathcal{V}_2$  belongs to  $V_1 \mathcal{V}_2 + \mathcal{V}_1 V_2$ . Then also the linearization of  $\mathcal{V}_1 \mathcal{V}_2$  equals this matrix subspace.

Geodesics and measuring curvature are closely related. Namely, the second fundamental form  $\Pi$  for the image of  $\Psi$  in a neighbourhood of  $(V_1, V_2)$  can be found by setting  $\frac{1}{2}(Q_{(V_1, V_2)}(W_1 + \tilde{W}_1, W_2 + \tilde{W}_2) - Q_{(V_1, V_2)}(W_1, W_2) - Q_{(V_1, V_2)}(\tilde{W}_1, \tilde{W}_2))$  to be

$$\Pi(W_1, W_2, \tilde{W}_1, \tilde{W}_2) = (I - \mathbf{P}_{V_1, V_2})(W_1 \tilde{W}_2 + \tilde{W}_1 W_2).$$

For more details on the second fundamental form and its geometric interpretation, see [26, p. 138].

As a first example, Theorem 2.6 corresponds to vanishing curvature. For another, with nonvanishing curvature, consider the subset of matrices related with approximating with the singular value decomposition.

**EXAMPLE 12.** This is Example 2 continued, i.e., consider  $\mathcal{F}_k \subset \mathbb{C}^{n \times n}$ , the set of matrices of rank  $k$  at most. Take  $(V_1, V_2)$  to be a generic point of  $\mathcal{V}_1 \times \mathcal{V}_2$  and let  $V_1 = U_1 \Sigma_1 W_1^*$  and  $V_2 = U_2 \Sigma_2 W_2^*$  be the singular value decompositions of  $V_1$  and  $V_2$ .<sup>13</sup> Then at  $\Psi(V_1, V_2)$  we have

$$V_1 \mathcal{V}_2 + \mathcal{V}_1 V_2 = U_1 \mathcal{W} W_2^*,$$

where  $\mathcal{W}$  is the matrix subspace of  $\mathbb{C}^{n \times n}$  consisting of matrices whose first  $k$  rows and columns can be freely chosen. Hence, the points  $(W_1, W_2) \in \mathcal{V}_1 \times \mathcal{V}_2$  satisfying  $Q_{(V_1, V_2)}(W_1, W_2) = 2W_1 W_2$  are readily determined. Consequently,  $\mathcal{V}_1 \mathcal{V}_2 = \mathcal{F}_k$  can be regarded as being maximally curved when measured in terms of (3.9). It is of interest to note that  $\mathcal{V}_2 \mathcal{V}_1$  is flat.

Observe that although  $Q_{(V_1, V_2)}$  yields a local measure of curvature, for the image of  $\Psi$  around the point  $\Psi(V_1, V_2)$ , it possesses a global character by the fact that the appearing orthogonal projector operates on the full image of  $\Psi$ . This provides an opportunity to use local information to draw conclusions about global properties of the set of product as follows.

**THEOREM 3.7.** *Let  $\mathcal{V}_1$  and  $\mathcal{V}_2$  be matrix subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$ . Then  $\overline{\mathcal{V}_1 \mathcal{V}_2}$  equals its linearization if and only if (3.8) vanishes for some  $(V_1, V_2)$ .*

<sup>13</sup>For  $j = 1, 2$ , the matrices  $U_j \in \mathbb{C}^{n \times n}$  and  $W_j \in \mathbb{C}^{n \times n}$  are unitary. In the matrices  $\Sigma_j$  only the first  $k$  diagonal entries can be nonzero. For more details on the singular value decomposition, see [13].

*Proof.* Assume  $\overline{\mathcal{V}_1\mathcal{V}_2}$  equals its linearization  $\mathcal{W}$ . By Sard's theorem, the set of critical values of  $\Psi$  is of the first category with Lebesgue measure zero; see, e.g., [30, p. 260, Theorem 1]. However, by Theorem 2.8, the image contains an open subset of  $\mathcal{W}$ . Its Lebesgue measure is obviously positive. Thereby there must be points in the image which are regular values. Consequently, (3.8) vanishes for some  $(V_1, V_2)$ .

For the converse, suppose (3.8) vanishes for some  $(V_1, V_2)$ . Clearly, then the closure of  $\mathcal{V}_1\mathcal{V}_2$  belongs to the respective matrix subspace  $\mathcal{W} = V_1\mathcal{V}_2 + \mathcal{V}_1V_2$ , i.e., the linearization of  $\mathcal{V}_1\mathcal{V}_2$ . We may regard  $\Psi : \mathcal{V}_1 \times \mathcal{V}_2 \rightarrow \mathcal{W}$ . By the constant rank theorem, the image of a neighbourhood of  $(V_1, V_2)$  under  $\Psi$  contains an open set (in the standard topology) in  $\mathcal{W}$ . In the Zariski topology, the image of  $\Psi$  is constructible and contains an open subset of its closure. This follows from a theorem of Chevalley. (For more details, see, e.g., [28, Theorem 10.2] or [17, p. 94].) Because the image of  $\Psi$  contains an open subset in the standard topology, the open subset of the image of  $\Psi$  in the Zariski topology is dense in  $\mathcal{W}$ .  $\square$

Obviously, under the assumptions of this theorem,  $\mathcal{W} = \overline{\mathcal{V}_1\mathcal{V}_2}$  is factorizable (assuming (3.2) is satisfied).

An immediate example is given by the LU factorization.

**EXAMPLE 13.** Using Theorem 3.7, it is straightforward to conclude that an LU factorization exists for elements in a dense open subset of  $\mathbb{C}^{n \times n}$ . (Of course, this is well known; a nonsingular matrix can be LU factored if and only if it is strongly nonsingular [19, p. 162].) Namely, let  $\mathcal{V}_1$  denote the set of lower triangular matrices and  $\mathcal{V}_2$  upper triangular matrices with constant diagonal. Then at  $(I, I) \in \mathcal{V}_1 \times \mathcal{V}_2$  the rank of  $\Psi$  is readily seen to be  $n^2$ . Consequently, (3.8) vanishes, so that  $\mathbb{C}^{n \times n} = \overline{\mathcal{V}_1\mathcal{V}_2}$ .

Of course, we may view this example just as a special case of the following consequence of Theorem 3.7.

**COROLLARY 3.8.** *Suppose  $\mathcal{V}_1$  and  $\mathcal{V}_2$  are matrix subspaces of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  both containing the identity. If  $\mathbb{C}^{n \times n} = \mathcal{V}_1 + \mathcal{V}_2$ , then  $\mathbb{C}^{n \times n} = \overline{\mathcal{V}_1\mathcal{V}_2}$ .*

Analogously can be formulated claims for lower (upper) triangular matrices in case of subspaces  $\mathcal{V}_1$  and  $\mathcal{V}_2$  of lower (upper) triangular matrices as follows.

**COROLLARY 3.9.** *Suppose  $\mathcal{W}$  is equivalent to a subalgebra of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  containing invertible elements. Then  $\mathcal{W}$  is factorizable.*

*Proof.* We may assume  $\mathcal{W}$  is a subalgebra of  $\mathbb{C}^{n \times n}$  over  $\mathbb{C}$  containing invertible elements. Observe that  $\mathcal{W}$  contains the identity. To see this, assume  $A \in \mathcal{W}$  is invertible. Then take a polynomial satisfying  $p(A) = A^{-1}$ . We have  $I = Ap(A) \in \mathcal{W}$ .

Denote by  $d$  the dimension of  $\mathcal{W}$ . Take a basis  $W_1, \dots, W_d$  of  $\mathcal{W}$  and set  $\mathcal{V}_1 = \text{span}\{I, W_1, W_2, \dots, W_k\}$  and  $\mathcal{V}_2 = \text{span}\{I, W_{k+1}, W_{k+2}, \dots, W_d\}$ . Clearly,  $\mathcal{V}_1\mathcal{V}_2 \subset \mathcal{W}$ . At  $(I, I)$  the rank of  $\Psi$  is  $d$ .  $\square$

For a classical factorization, any matrix  $A \in \mathbb{C}^{n \times n}$  is the product of two symmetric matrices such that there are at least  $n$  degrees of freedom to construct a factorization [20]. (See also Example 8.) This redundancy can be reduced as follows.

**PROPOSITION 3.10.** *Let  $\mathcal{V}_1$  be the set of symmetric matrices and  $\mathcal{V}_2$  the subset of symmetric matrices having constant antidiagonal. Then  $\mathbb{C}^{n \times n} = \overline{\mathcal{V}_1\mathcal{V}_2}$ .*

*Proof.* Suppose  $n$  is odd and denote by  $J$  the permutation matrix having ones on its antidiagonal. Then, for a generic diagonal matrix  $D = \text{diag}(d_1, d_2, \dots, d_n)$ , the rank of  $\Psi$  at  $(D, J) \in \mathcal{V}_1 \times \mathcal{V}_2$  can be shown to equal  $n^2$ . (For  $n$  even, proceed similarly with  $(J, D) \in \mathcal{V}_1 \times \mathcal{V}_2$ .) To see this, consider  $DV_2 + V_1J$ . Clearly, the  $(j, k)$  entry of  $V_1J$  is the  $(j, n - k + 1)$  entry of  $V_1$ . For  $k \neq j$  and  $k \neq n - j + 1$  this means that the entries in  $DV_2 + V_1J$  located symmetrically with respect to the diagonal and antidiagonal are interdependent, i.e., the  $(j, k)$ ,  $(k, j)$ ,  $(n - k + 1, n - j + 1)$  and

$(n-j+1, n-k+1)$  entries. To satisfy  $D\mathcal{V}_2 + \mathcal{V}_1J = \mathbb{C}^{n \times n}$  yields us the linear system

$$\begin{bmatrix} d_j & 0 & 1 & 0 \\ d_k & 0 & 0 & 1 \\ 0 & d_{n-k+1} & 1 & 0 \\ 0 & d_{n-j+1} & 0 & 1 \end{bmatrix} \begin{bmatrix} V_2(j,k) \\ V_2(n-k+1, n-j+1) \\ V_1(j, n-k+1) \\ V_1(k, n-j+1) \end{bmatrix} = \begin{bmatrix} M(j,k) \\ M(k,j) \\ M(n-k+1, n-j+1) \\ M(n-j+1, n-k+1) \end{bmatrix}$$

which should have a solution for any right-hand side. This is possible if and only if  $1 - \frac{d_k d_{n-k+1}}{d_j d_{n-j+1}} \neq 0$  for  $j \neq k$  and  $k \neq n-j+1$ . For  $k = j$  and  $k = n-j+1$  we obtain the condition  $d_j \neq 0$ . These are the genericity conditions the matrix  $D$  needs to satisfy.  $\square$

Since  $\mathcal{V}_1$  is an invertible matrix subspace, the factorization of the proposition can be computed with the methods proposed in [20].

**4. Conclusions.** Motivated by factorization problems, the set of products of two matrix subspaces is studied. Differential geometric approach applied to the respective smooth mapping yields a measure of curvature for the set of products. Its vanishing corresponds to the concept of factorizable matrix subspace. The notion of irreducible matrix subspace was introduced. The LU factorization, the singular value decomposition and the Craig-Sakamoto theorem served as illustrative examples how seemingly different concepts (algorithmically at least) can be put under the same caption.

**Appendix: Hilbert Nullstellensatz.** Let  $p_1, \dots, p_k$  be complex polynomials in  $n$  variables. There is no common zero in  $\mathbb{C}^n$  if and only if there are complex polynomials  $q_j$  in  $n$  variables satisfying

$$\sum_{j=1}^k q_j p_j = 1.$$

This is the Hilbert Nullstellensatz. The effective Hilbert Nullstellensatz states that the degrees of the  $q_j$  may be assumed to satisfy

$$\deg q_j \leq \begin{cases} D^n & \text{if } D \geq 3 \\ 2^{\min\{n,k\}} & \text{if } D = 2 \end{cases},$$

where  $D = \max \deg p_j$ . See [2], [23] and [31].

#### REFERENCES

- [1] M. BRIN, *On the Zappa-Szép product*, Comm. Algebra, 33 (2005), pp. 393–424.
- [2] W. BROWNAWELL, *Bounds for the degrees in the Nullstellensatz*, Ann. Math., 126 (1987), pp. 577–591.
- [3] J. BUSS, G. FRANSDEN AND J. SHALLIT, *The computational complexity of some problems of linear algebra*, J. Comput. System Sci., 58 (1999), pp. 572–596.
- [4] M. BYCKLING AND M. HUHTANEN, *Approximate factoring of the inverse*, Numer. Math., 117 (2011), pp. 507–528.
- [5] L. CONLON, *Differential Manifolds, a First Course*, Birkhäuser, Boston, 1993.
- [6] D. COX, J. LITTLE AND D. O’SHEA, *Ideals, Varieties and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Springer-Verlag, New York, 1997.
- [7] C.W. CURTIS AND I. REINER, *Representation Theory of Finite Groups and Associative Algebras*, AMS Chelsea Publishing, 1962.
- [8] M. DUMAIS AND G.P. STYAN, *A bibliography on the distribution of quadratic forms in normal variables, with special emphasis on the Craig-Sakamoto theorem and on Cochran’s theorem*, In G. Styan, ed., *Three Bibliographies and a Guide*, Seventh International Workshop on Matrices and Statistics Fort Lauderdale, 1998, pp. 1–9.

- [9] E.G. EFFROS, Z.-J. RUAN, *Operator Spaces*, Oxford University Press, Oxford, 2000.
- [10] J.-C. FAUGÈRE, M. EL DIN AND P.-J. SPAENLEHAUER, *Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1,1): algorithms and complexity*, J. Sym. Comp., 46 (2011), pp. 406–437.
- [11] J.-C. FAUGÈRE, M. EL DIN AND P.-J. SPAENLEHAUER, *Computing loci of rank defects of linear matrices using Gröbner bases and applications to cryptology*, In ISSAC '10: Proceedings of the 2010 international symposium on Symbolic and algebraic computation, ISSAC '10, pp. 257–264, New York, NY, USA, 2010.
- [12] I. GOHBERG AND V. OLSHEVSKY, *Circulants, displacements and decompositions of matrices*, Integr. Equat. Oper. Th., 15 (1992), pp. 730–743.
- [13] G.H. GOLUB AND C.F. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, the 3rd ed., 1996.
- [14] M.L. GREEN, *Holomorphic maps into complex projective space omitting hyperplanes*, Trans. Amer. Math. Soc., 169 (1972), pp. 89–103.
- [15] P. HALMOS, *Bad products of good matrices*, Linear and Multilinear Algebra, 29 (1991), no. 1, pp. 1–20.
- [16] J. HARRIS, *Algebraic Geometry, A First Course*, Springer, New York, 1992.
- [17] R. HARTSHORNE, *Algebraic Geometry*, Springer-Verlag, New York, Heidelberg, Berlin, 1977.
- [18] O. HOLTZ, V. MEHRMANN AND H. SCHNEIDER, *Potter, Wielandt, and Drazin on the matrix equation  $AB = \omega BA$ : new answers to old questions*, Amer. Math. Monthly, 111 (2004), no. 8, pp. 655–667.
- [19] R.A. HORN AND C.R. JOHNSON, *Topics in Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1991.
- [20] M. HUHTANEN, *Factoring matrices into the product of two matrices*, BIT, 47 (2007), pp. 793–808.
- [21] M. HUHTANEN, *Matrix subspaces and determinantal hypersurfaces*, Ark. Mat., 48 (2010), pp. 57–77.
- [22] M. HUHTANEN, *Differential geometry of matrix inversion*, Math. Scand., 107 (2010), pp. 267–284.
- [23] J. KOLLÁR, *Sharp effective Nullstellensatz*, J. Amer. Math. Soc., 1 (1988), pp. 963–975.
- [24] T.J. LAFFEY, *Conjugacy and factorization results on matrix groups*, Polish Academy of Sciences, Warsaw (eds). Functional Analysis and Operator Theory, Banach Center Publications Number 30, (1994), pp. 203–221.
- [25] P. LAX, *On the factorization of matrix valued functions*, Comm. Pure Appl. Math., 24 (1976), pp. 683–688.
- [26] J.M. LEE, *Riemannian Manifolds, An Introduction to Curvature*, Springer Verlag, New York, 1997.
- [27] J.M. LEE, *Introduction to Topological Manifolds*, Springer Verlag, New York, 2000.
- [28] J.S. MILNE, *Algebraic Geometry, Lecture notes at [www.jmilne.org/math/](http://www.jmilne.org/math/)*.
- [29] V. PAULSEN, *Completely Bounded Maps and Operator Algebras*, Cambridge University Press, Cambridge, 2002.
- [30] M. POSTNIKOV, *Smooth manifolds (Lectures in geometry)*, Mir Publishers, Moscow, 1989.
- [31] M. SOMBRA, *A sparse effective Nullstellensatz*, Adv. Appl. Math., 22 (1999), pp. 271–295.
- [32] O. TAUSSKY, *The characteristic polynomial and the characteristic curve of a pencil of matrices with complex entries*, Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber., II 195 (1986), no. 1-3, pp. 175–178.
- [33] P.Y. WU, *The operator factorization problems*, Linear Algebra Appl., 117 (1989), pp. 35–63.